

Figure 1 125P5C8 SSH sequence

GATCACGTGCTGTCGATATCCTTCACATTGCCATGTTTCAGTGAGCTGTAGATAATCTCTGGAGCCAGGTGCTGAAGTGA
TATATCCCAGAAATATCACTTGATTAGAGCTACTTTTCAGTAGTTTTGAAACAGCAATAGCCTGCAGTTTCCTGTCGAG
GTCATCTTCGTGGTTCCCAAAGTGTGTACGACAAAATCCACCAGCTTGCCCCGAAATGTTAACGGTCAATGTGATGGCT
GGTGCGATCTTGCTGTGTTGGCCAGGCTGGTCTCAACGTGCAGATAGATC

125P5C8 SSH sequence

Figure 2 Double stranded sequence and ORF for 125P5C8-Pro-pCR2.1.

	M	T	S	L	W	R	E	I	L	L	E	S	L	L	G	C	V	S
1	ATG	ACC	TCG	CTG	TGG	AGA	GAA	ATC	CTC	TTG	GAG	TCG	CTG	CTG	GGA	TGT	GTT	TCT
	W	S	L	Y	H	D	L	G	P	M	I	Y	Y	F	P	L	Q	T
55	TGG	TCT	CTC	TAC	CAT	GAC	CTG	GGA	CCG	ATG	ATC	TAT	TAC	TTT	CCT	TTG	CAA	ACA
	L	E	L	T	G	L	E	G	F	S	I	A	F	L	S	P	I	F
109	CTA	GAA	CTC	ACT	GGG	CTT	GAA	GGT	TTT	AGT	ATA	GCA	TTT	CTT	TCT	CCA	ATA	TTC
	L	T	I	T	P	F	W	K	L	V	N	K	K	W	M	L	T	L
163	CTA	ACA	ATT	ACT	CCT	TTC	TGG	AAA	TTG	GTT	AAC	AAG	AAG	TGG	ATG	CTA	ACC	CTG
	L	R	I	I	T	I	G	S	I	A	S	F	Q	A	P	N	A	K
217	CTG	AGG	ATA	ATC	ACT	ATT	GGC	AGC	ATA	GCC	TCC	TTC	CAG	GCT	CCA	AAT	GCC	AAA
	L	R	L	M	V	L	A	L	G	V	S	S	S	L	I	V	Q	A
271	CTT	CGA	CTG	ATG	GTT	CTT	GCG	CTT	GGG	GTG	TCT	TCC	TCA	CTG	ATA	GTG	CAA	GCT
	V	T	W	W	S	G	S	H	L	Q	R	Y	L	R	I	W	G	F
325	GTG	ACT	TGG	TGG	TCG	GGA	AGT	CAT	TTG	CAA	AGG	TAC	CTC	AGA	ATT	TGG	GGA	TTC
	I	L	G	Q	I	V	L	V	V	L	R	I	W	Y	T	S	L	N
379	ATT	TTA	GGA	CAG	ATT	GTT	CTT	GTT	CTA	CGC	ATA	TGG	TAT	ACT	TCA	CTA	AAC	
	P	I	W	S	Y	Q	M	S	N	K	V	I	L	T	L	S	A	I
433	CCA	ATC	TGG	AGT	TAT	CAG	ATG	TCC	AAC	AAA	GTG	ATA	CTG	ACA	TTA	AGT	GCC	ATA
	A	T	L	D	R	I	G	T	D	G	D	C	S	K	P	E	E	K
487	GCC	ACA	CTT	GAT	CGT	ATT	GGC	ACA	GAT	GGT	GAC	TGC	AGT	AAA	CCT	GAA	GAA	AAG
	K	T	G	E	V	A	T	G	M	A	S	R	P	N	W	L	L	A
541	AAG	ACT	GGT	GAG	GTA	GCC	ACG	GGG	ATG	GCC	TCT	AGA	CCC	AAC	TGG	CTG	CTG	GCA
	G	A	A	F	G	S	L	V	F	L	T	H	W	V	F	G	E	V
595	GGG	GCT	GCT	TTT	GGT	AGC	CTT	GTG	TTC	CTC	ACC	CAC	TGG	GTT	TTT	GGA	GAA	GTC
	S	L	V	S	R	W	A	V	S	G	H	P	H	P	G	P	D	P
649	TCT	CTT	GTT	TCC	AGA	TGG	GCA	GTG	AGT	GGG	CAT	CCA	CAT	CCA	GGG	CCA	GAT	CCT
	N	P	F	G	G	A	V	L	L	C	L	A	S	G	L	M	L	P
703	AAC	CCA	TTT	GGA	GGT	GCA	GTA	CTG	CTG	TGC	TTG	GCA	AGT	GGA	TTG	ATG	CTT	CCA
	S	C	L	W	F	R	G	T	G	L	I	W	W	V	T	G	T	A
757	TCT	TGT	TTG	TGG	TTT	CGT	GGT	ACT	GGT	TTG	ATC	TGG	TGG	GTT	ACA	GGA	ACA	GCT
	S	A	A	G	L	L	Y	L	H	T	W	A	A	A	V	S	G	C
811	TCA	GCT	GCG	GGG	CTC	CTT	TAC	CTG	CAC	ACA	TGG	GCA	GCT	GCT	GTG	TCT	GGC	TGT
	V	F	A	I	F	T	A	S	M	W	P	Q	T	L	G	H	L	I
865	GTC	TTC	GCC	ATC	TTT	ACT	GCA	TCC	ATG	TGG	CCC	CAA	ACA	CTT	GGA	CAC	CTT	ATT
	N	S	G	T	N	P	G	K	T	M	T	I	A	M	I	F	Y	L
919	AAC	TCA	GGG	ACA	AAC	CCT	GGG	AAA	ACC	ATG	ACC	ATT	GCC	ATG	ATA	TTT	TAT	CTT
	L	E	I	F	F	C	A	W	C	T	A	F	K	F	V	P	G	G
973	CTA	GAA	ATA	TTT	TTC	TGT	GCC	TGG	TGC	ACA	GCT	TTT	AAG	TTT	GTC	CCA	GGA	GGT
	V	Y	A	R	E	R	S	D	V	L	L	G	T	M	M	L	I	I
1027	GTC	TAC	GCT	AGA	GAA	AGA	TCA	GAT	GTG	CTT	TTG	GGG	ACA	ATG	ATG	TTA	ATT	ATC
	G	L	N	M	L	F	G	P	K	K	N	L	D	L	L	L	Q	T
1081	GGG	CTG	AAT	ATG	CTA	TTT	GGT	CCT	AAG	AAA	AAC	CTT	GAC	TTG	CTT	CTT	CAA	ACA
	K	N	S	S	K	V	L	F	R	K	S	E	K	Y	M	K	L	F
1135	AAA	AAC	AGT	TCT	AAA	GTG	CTT	TTC	AGA	AAG	AGT	GAA	AAA	TAC	ATG	AAA	CTT	TTT
	L	W	L	L	V	G	V	G	L	L	G	L	G	L	R	H	K	A
1189	CTG	TGG	CTG	CTT	GTT	GGT	GTG	GGA	TTG	TTG	GGA	TTA	GGA	CTA	CGG	CAT	AAA	GCC
	Y	E	R	K	L	G	K	V	A	P	T	K	E	V	S	A	A	I
1243	TAT	GAG	AGA	AAA	CTG	GGC	AAA	GTG	GCA	CCA	ACC	AAA	GAG	GTC	TCT	GCT	GCC	ATC
	W	P	F	R	F	G	Y	D	N	E	G	W	S	S	L	E	R	S
1297	TGG	CCT	TTC	AGG	TTT	GGA	TAT	GAC	AAT	GAA	GGG	TGG	TCT	AGT	CTA	GAA	AGA	TCA
	A	H	L	L	N	E	T	G	A	D	F	I	T	I	L	E	S	D
1351	GCT	CAC	CTG	CTC	AAT	GAA	ACA	GGT	GCA	GAT	TTC	ATA	ACA	ATT	TTG	GAG	AGT	GAT
	A	S	K	P	Y	M	G	N	N	D	L	T	M	W	L	G	E	K
1405	GCT	TCT	AAG	CCC	TAT	ATG	GGG	AAC	AAT	GAC	TTA	ACC	ATG	TGG	CTA	GGG	GAA	AAG
	L	G	F	Y	T	D	F	G	P	S	T	R	Y	H	T	W	G	I
1459	TTG	GGT	TTC	TAT	ACA	GAC	TTT	GGT	CCA	AGC	ACA	AGG	TAT	CAC	ACT	TGG	GGG	ATT
	M	A	L	S	R	Y	P	I	V	K	S	E	H	H	L	L	P	S
1513	ATG	GCT	TTG	TCA	AGA	TAC	CCA	ATT	GTG	AAA	TCT	GAG	CAT	CAC	CTT	CTT	CCG	TCA
	P	E	G	E	I	A	P	A	I	T	L	T	V	N	I	S	G	K
1567	CCA	GAG	GGC	GAG	ATC	GCA	CCA	GCC	ATC	ACA	TTG	ACC	GTT	AAC	ATT	TCG	GGC	AAG
	L	V	D	F	V	V	T	H	F	G	N	H	E	D	D	L	D	R

Figure 2 Double stranded sequence and ORF for 125P5C8-Pro-pCR2.1.

Figure 2 Cont'd

1621	CTG GTG GAT TTT GTC GTG ACA CAC TTT GGG AAC CAC GAA GAT GAC CTC GAC AGG
	K L Q A I A V S K L L K S S S N Q V
1675	AAA CTG CAG GCT ATT GCT GTT TCA AAA CTA CTG AAA AGT AGC TCT AAT CAA GTG
	I F L G Y I T S A P G S R D Y L Q L
1729	ATA TTT CTG GGA TAT ATC ACT TCA GCA CCT GGC TCC AGA GAT TAT CTA CAG CTC
	T E H G N V K D I D S T D H D R W C
1783	ACT GAA CAT GGC AAT GTG AAG GAT ATC GAC AGC ACT GAT CAT GAC AGA TGG TGT
	E Y I M Y R G L I R L G Y A R I S H
1837	GAA TAC ATT ATG TAT CGA GGG CTG ATC AGG TTG GGT TAT GCA AGA ATC TCC CAT
	A E L S D S E I Q M A K F R I P D D
1891	GCT GAA CTG AGT GAT TCA GAA ATT CAG ATG GCA AAA TTT AGG ATC CCT GAT GAC
	P T N Y R D N Q K V V I D H R E V S
1945	CCC ACT AAT TAT AGA GAC AAC CAG AAA GTG GTC ATA GAC CAC AGA GAA GTT TCT
	E K I H F N P R F G S Y K E G H N Y
1999	GAG AAA ATT CAT TTT AAT CCC AGA TTT GGA TCC TAC AAA GAA GGA CAC AAT TAT
	E N N H H F H M N T P K Y F L *
2053	GAA AAC AAC CAT CAT TTT CAT ATG AAT ACT CCC AAA TAC TTT TTA TGA AAC

1621 1675 1729 1783 1837 1891 1945 1999 2053

Figure 3 Amino acid sequence of 125P5C8.

1	MTSLWREILL	ESLLGCVSWS	LYHDLGPMIY	YFPLQTLELT	GLEGFSIAFL	50
51	SPIFLTITPF	WKLVNKKWML	TLLRIITIGS	IASFQAPNAK	LRLMVLALGV	100
101	SSSLIVQAVT	WWSGSHLQRY	LRIWGFILGQ	IVLVVLRIWY	TSLNPIWSYQ	150
151	MSNKVILTLS	AIATLDRIGT	DGDCSKPEEK	KTGEVATGMA	SRPNWLLAGA	200
201	AFGSLVFLTH	WVFGEVSLVS	RWAVSGHPPH	GPDPNPFGGA	VLLCLASGLM	250
251	LPSCLWFRGT	GLIWVVTGTA	SAAGLLYLHT	WAAAVSGCVF	AIFTASMWPQ	300
301	TLGHLINSGT	NPGKTMTIAM	IFYLLEIFFC	AWCTAFKFVP	GGVYARERSD	350
351	VLLGTMMMLII	GLNMLFGPKK	NLDLLLQTKN	SSKVLFRKSE	KYMKLFLWLL	400
401	VGVGLLGLGL	RHKAYERKLG	KVAPTKEVSA	AIWPFRFGYD	NEGWSSLERS	450
451	AHLLNETGAD	FITILESDAS	KPYMGNNDLT	MWLGEKLGFY	TDFGPSTRYH	500
501	TWGIMALSRY	PIVKSEHLL	PSPEGEIAPA	ITLTVNISGK	LVDFVVTTHFG	550
551	NHEDDLDRKL	QAIAVSKLLK	SSSNQVIFLG	YITSAPGSRD	YLQLTEHGNV	600
601	KDIDSTDHDR	WCEYIMYRGL	IRLGYARISH	AELSDSEIQM	AKFRIPDDPT	650
651	NYRDNQKVI	DHREVSEKIH	FNPRFGSYKE	GHNYENNHHF	HMNTPKYFL	699

125P5C8

Figure 4A Alignment with AK025164 protein product

Score = 1397 bits (3615), Expect = 0.0

Identities = 682/699 (97%), Positives = 683/699 (97%)

```

Query: 1  MTSLWREILLESLLGCVSWSLYHDLGPMIYFPLQTLELTGLEGFSIAFLSPIFLTITPF 60
          MTSLWREILLESLLGCVSWSLYHDLGPMIYFPLQTLELTGLEGFSIAFLSPIFLTITPF
Sbjct: 1  MTSLWREILLESLLGCVSWSLYHDLGPMIYFPLQTLELTGLEGFSIAFLSPIFLTITPF 60

Query: 61  WKLVNKKWMLTLLRIITIGSIASFQAPNAKLRLMVLALGVSSSLIVQAVTWWSGSHLQRY 120
          WKLVNKKWMLTLLRIITIGSIASFQAPNAKLRLMVLALGVSSSLIVQAVTWWSGSHLQRY
Sbjct: 61  WKLVNKKWMLTLLRIITIGSIASFQAPNAKLRLMVLALGVSSSLIVQAVTWWSGSHLQRY 120

Query: 121  LRIWGFILGQIVLVVLRIRIWTSLNPIWSYQMSNKVILTLSAIATLDRIGTDGDCSKPEEK 180
          LRIWGFILGQIVLVVLRIRIWTSLNPIWSYQMSNKVILTLSAIATLDRIGTDGDCSKPEEK
Sbjct: 121  LRIWGFILGQIVLVVLRIRIWTSLNPIWSYQMSNKVILTLSAIATLDRIGTDGDCSKPEEK 180

Query: 181  KTGEVATGMASRPNWLLAGAAFGSLVFLTHWVFGEVSLVSRWAVSGHHPGPDNPFGGA 240
          KTGEVATGMASRPNWLLAGAAFGSLVFLTHWVFGEVSLVSRWAVSGHHPGPDNPFGGA
Sbjct: 181  KTGEVATGMASRPNWLLAGAAFGSLVFLTHWVFGEVSLVSRWAVSGHHPGPDNPFGGA 240

Query: 241  VLLCLASGLMLPSCLWFRGTGLIWWVTGTASAAGLLYLHTWAAAVSGCVFAIFTASMWPQ 300
          VLLCLASGLMLPSCLWFRGTGLIWWVTGTASAAGLLYLHTWAAAVSGCVFAIFTASMWPQ
Sbjct: 241  VLLCLASGLMLPSCLWFRGTGLIWWVTGTASAAGLLYLHTWAAAVSGCVFAIFTASMWPQ 300

Query: 301  TLGHLINSGTNPCKTMTIAMIFYLLEIFFCAWCTAFKFVPGGVYARERSDVLLGTMMMLII 360
          TLGHLINSGTNPCKTMTIAMIFYLLEIFFCAWCTAFKFVPGGVYARERSDVLLGTMMMLII
Sbjct: 301  TLGHLINSGTNPCKTMTIAMIFYLLEIFFCAWCTAFKFVPGGVYARERSDVLLGTMMMLII 360

Query: 361  GLNMLFGPKKNLDLLLQTKNSSKVLFRKSEKYMXXXXXXXXXXXXXXXXXXXXXRHKAYERKLG 420
          GLNMLFGPKKNLDLLLQTKNSSKVLFRKSEKYM                                RHKAYERKLG
Sbjct: 361  GLNMLFGPKKNLDLLLQTKNSSKVLFRKSEKYMKFLWLLVGVLGLGLRHKAYERKLG 420

Query: 421  KVAPTKEVSAAIWPFRFGYDNEGWSLERSAHLNETGADFITILES DASKPYMGNNDLT 480
          KVAPTKEVSAAIWPFRFGYDNEGWSLERSAHLNETGADFITILES DASKPYMGNNDLT
Sbjct: 421  KVAPTKEVSAAIWPFRFGYDNEGWSLERSAHLNETGADFITILES DASKPYMGNNDLT 480

Query: 481  MWLGEKLGFTYDFGPSTRYHTWGIMALSRYP IVKSEHHLLPSPEGEIAPAITLTVNISGK 540
          MWLGEKLGFTYDFGPSTRYHTWGIMALSRYP IVKSEHHLLPSPEGEIAPAITLTVNISGK
Sbjct: 481  MWLGEKLGFTYDFGPSTRYHTWGIMALSRYP IVKSEHHLLPSPEGEIAPAITLTVNISGK 540

Query: 541  LVDFVVTHTFGNHEDDLDRKLQAI AVSKLLKSSSNQVIFLGYITSAPGSRDYLQLTEHGNV 600
          LVDFVVTHTFGNHEDDLDRKLQAI AVSKLLKSSSNQVIFLGYITSAPGSRDYLQLTEHGNV
Sbjct: 541  LVDFVVTHTFGNHEDDLDRKLQAI AVSKLLKSSSNQVIFLGYITSAPGSRDYLQLTEHGNV 600

Query: 601  KDIDSTDHWRWCEYIMYRGLIRLGYARISHAELSDSEIQMAKFRIPDDPTNYRDNQKVVI 660
          KDIDSTDHWRWCEYIMYRGLIRLGYARISHAELSDSEIQMAKFRIPDDPTNYRDNQKVVI
Sbjct: 601  KDIDSTDHWRWCEYIMYRGLIRLGYARISHAELSDSEIQMAKFRIPDDPTNYRDNQKVVI 660

Query: 661  DHREVSEKIHFNPRFGSYKEGHNYENNHHFHMNTPKYFL 699
          DHREVSEKIHFNPRFGSYKEGHNYENNHHFHMNTPKYFL
Sbjct: 661  DHREVSEKIHFNPRFGSYKEGHNYENNHHFHMNTPKYFL 699

```

Figure 4B Alignment with yeast YCR017 cp Protein

Score = 261 bits (668), Expect = 1e-68

Identities = 204/705 (28%), Positives = 330/705 (45%), Gaps = 50/705 (7%)

Query: 15 GCVSWSLYHDLGPMIYYFPLQTLELTGLEGFSAIFLSPIFLTITPFWKLVNKKWMLTLLR 74
G + WS L I++FPL + ++G E + +L PIFL + PF ++ + L

Sbjct: 279 GFLFWSNVTSLLCSIWHFPLWYMGISGYEAAILGYLGPIFLYL-PFVSEAFQTQYGVLLGG 337

Query: 75 IITIGSIASFQAPNAKLRLMVLALGVSSSLIVQAVTWWSGSHLQ-RYLRIWGFILGQIVL 133
II IG+ Q P +L + + ++ + VQ + + + + + W +LG +

Sbjct: 338 IIAIGAYI-VQMPELRLISVAVGTSITVATFVQNLRYITNAETSFSFALTW--LLGLVAS 394

Query: 134 VVLRIWYTSLNPIWSYQMS-----NKVILTLSAIATLDRIGTDGDCSKPEEKKTGEVATG 188
V+L++ + + NP W NK L L+ + + + + E K+ + +

Sbjct: 395 VILKMGFYTNNTPTWVILDERNGGYNKLTALVLTVLFGM--LSPYVNSINFEGKRNAQAKS- 451

Query: 189 MASRPNWLLAGAAFGSLVFLTHWVFGEVSLVSRWAVSGHPPH-GPDPNPFGGAVLLCLAS 247
AS L FGSL+F H + + S WA G+ GP P P+G L C

Sbjct: 452 -ASLIGKLFLAVGFGSLLFGIHQLLTDSSSTTIYWAWEGYNESHGPLPWPWGA--LTCTVM 508

Query: 248 GLMLPSCWFRGTGLIWWVTGTASAAGLLY--LHTWAAAV-SGCVFAIFTASMWPQ---T 301
S + F G L+ + S A L + W + G ++AI + P

Sbjct: 509 LFASLSSVKFMGKPLVPCLLLLISTAVLSARSITQWPKYIFGGLLYAIAMLWLVPYSYFSA 568

Query: 302 LGHLINSGTNPCKTMTIAMIFYLLEIFFCAWCTAFKFVPGGVYARERSDVLLGTMMLIIG 361
LG + N ++ Y++ + W A+ FVP G RE+ + +L I

Sbjct: 569 LGQVQNIWV-----YVLSFSVYIIFVLAHVWVYAFVPMGWVLREKIETVLAFSSTFII 623

Query: 362 LNMLFGPKKNLDLLLQTKNSSKVLFRKSEKYMKXXXXXXXXXXXXXXXXXXXXRHKAYERKLGK 421
+ L N+ L+ K +F R R G

Sbjct: 624 IGALTCKNLNIQLVTMGKKFFIYVF-----FFAVALLSLTARFVYDIRPTGI 670

Query: 422 VAP----TKEVSAAIWPFRTGYDNEGWSLERSAHLNETGADFITILES DASKPYMGNN 477
P ++ ++A IW FG DN+ W+S +R +L+ + D + +LE+D + MGN

Sbjct: 671 PQPYHPDSQLITAGIWTIHFGLDNDMWASEDRMINLIKDMELDVVGLLETDTQIRITMGNR 730

Query: 478 DLTMWLGEKLGIFYTDFGPSTRYHTWGIMALSRYPIVKSEHLLPSPEGEIAPAITLTV-N 536
DLT L L Y DFGP HTWG + LS++PIV S HHLLPSP GE+APAI T+

Sbjct: 731 DLTSKLAHDLNMYADFGPGPNKHTWGCVLSSKFPVIVNSTHLLPSPVGELAPAIHATLQT 790

Query: 537 ISGKLVDVFVTHFGNHEDDLDRKLQAIIVSKLLKSSSNQVIFLGYITSAPGSRDY-LQLT 595
+ LVD V H G ED+ DR+LQ+ ++KL+ +++ I L Y+ PG +Y ++

Sbjct: 791 YNDTLVDVFVHSGQEEDRRLQSNYMAKLMGNTTRPAILLSYLVDPGEGNYNTYVS 850

Query: 596 EHGNVKDIDSTDHWRCEYIMYRGLIRLGYARISHAELSDSEIQMAKFRIPDDPTNYRDN 655
E + DID +D DRWCEYI+YRGL R GYAR++ ++D+E+Q+ KF++ + ++

Sbjct: 851 ETSGMHDIDPSDDDRWCEYIILYRGLRRTGYARVARGTITDTELQVGKFQVLSEQA-LVEH 909

Query: 656 QKVVIDHREVSEKIHFNPRFGSYKEGHNYENNHHFHM-NTPKYFL 699
+ ++ +SE + + +F G E H +H+ + P+Y+L

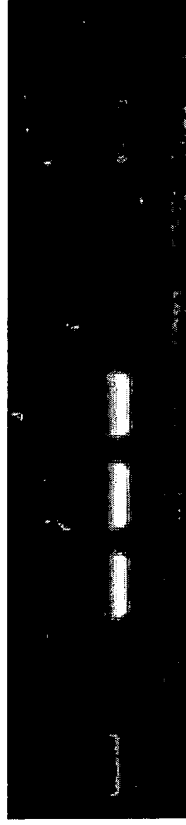
Sbjct: 910 SDSMYEYGHMSEPEYEDMKFPDKFLGEG-ERGHFYHVFDEPRYYL 953

Figure 5

VP1 VP2 XeP N.Pr PrCa BICa KiCa CoCa H2O



26X



30X

Figure 6A

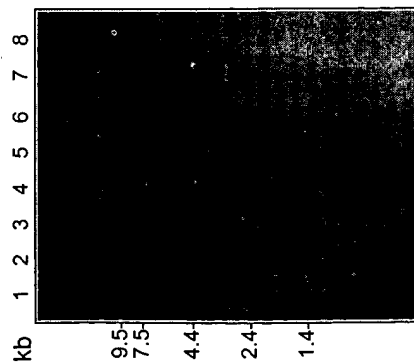


Figure 6B

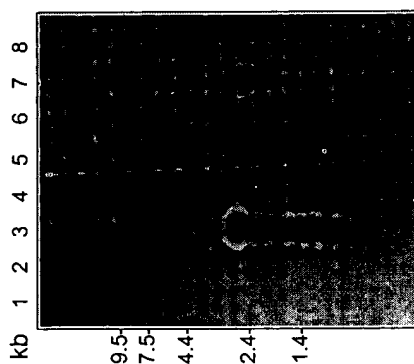


Figure 6C

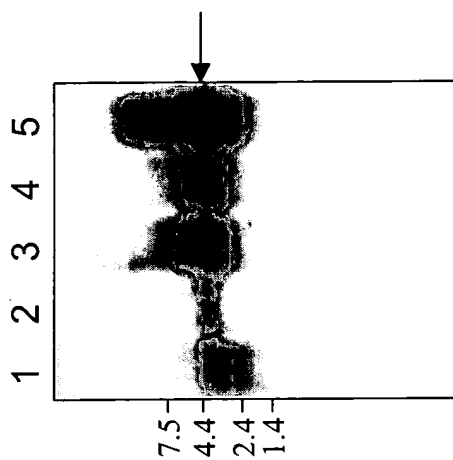


Figure 7

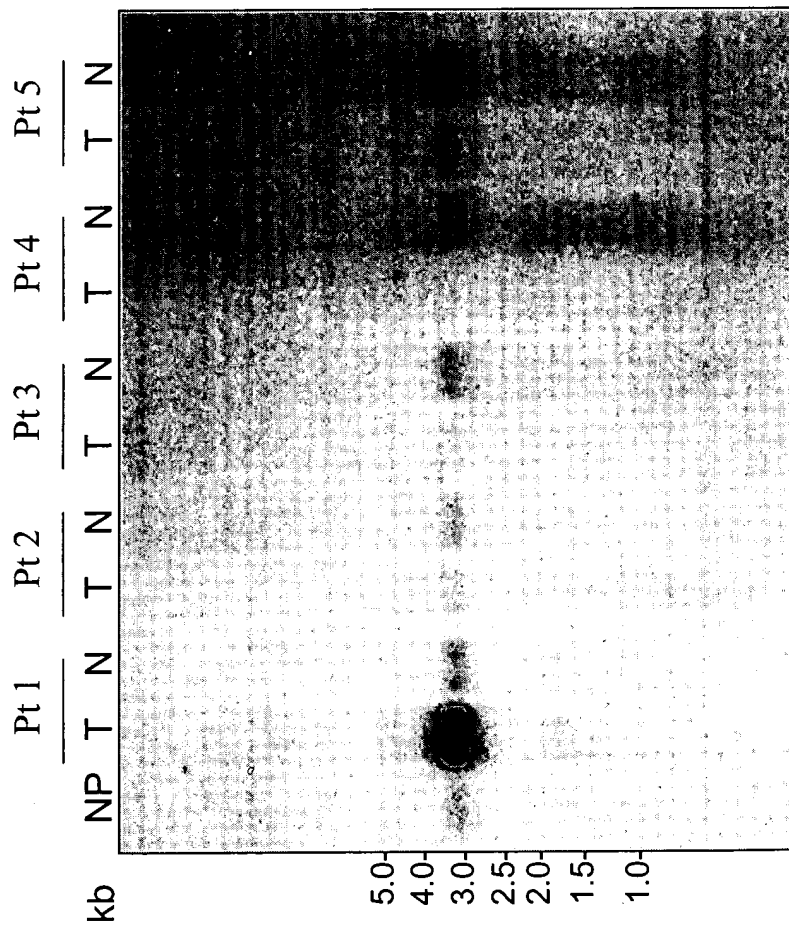


Figure 8

